

## DSP16210 을 이용한 8 kbps CS-ACELP 의 실시간 구현

\* 박지현, 박성일, 정원국, 임병근

† LG 정보통신(주) 중앙연구소

### Real-Time Implementation of the 8 kbps CS-ACELP Using DSP16210 Fixed Point DSP Chip

\*Jee Hyun Park, Sung Il Park, Weon Gook Chung, Byoung Keun Lim

†R&D Center, LG Information & Communications, Ltd.

Telephone: +82-343-450-7877

Fax: +82-343-450-7050

E-mail: jeehp@lgic.co.kr

**Abstract** - Real-time implementation of Conjugate-Structure Algebraic CELP(CS-ACELP) is presented. ITU-T Study Group(SG) 15 has standardized the CS-ACELP speech coding algorithm as G.729. A real-time implementation of the CS-ACELP is achieved using 16 bit fixed point DSP16210 Digital Signal Processor (DSP) of Lucent Technologies. The speech coder has been implemented in the bit-exact manner using the fixed point CS-ACELP C source which is the part of the G.729 standard. To provide a multi-channel vocoder solution to digital communication system, we try to minimize the complexity(e.g., MIPS, ROM, RAM) of CS-ACELP. Our speech coder shows 15.5 MIPS in performance which enables 4 channel CS-ACELP to be processed with one DSP16210.

### 1. 서 론

디지털 이동통신 시스템에서 음성 부호화기는 통화 품질에 직접적으로 영향을 주는 부분중의 하나로서 낮은 전송률과 무선환경에 대한 강인성이 있으며 유선전화의 음질(toll quality)을 가져야 한다. 최근에 ITU-T에서 차세대 통신망(IMT-2000)과 개인 휴대 통신(PCS)에 응용하기 위해 표준안으로 채택한 Conjugate Structure Algebraic CELP (CS-ACELP, G.729)[1][2]는 8 kbps의 전송률과 채널 에러에 강한 특성을 가지며 잡음/비잡음 환경에서 32 kbps ADPCM(G.726)과 대등한 음질을 나타내는 것으로 알려져 있다[2-4].

CS-ACELP 음성 부호화 알고리즘은 CELP 구조를 바탕으로 피치 및 코드북(codebook)의 이득을 양자화하는

데 결합구조(conjugate structure)[5] 형태를 가지는 2개의 벡터 테이블을 이용하고 여기 신호로는 대수적(algebraic)[6] 코드북을 사용하는 특징을 가진다.

최근에 개발된 신호처리 프로세서(DSP)들은 과거에 비해 향상된 성능과 많은 내부 메모리를 가지고 있어서 다수 채널을 서비스할 수 있는 음성 부호화기를 하나의 DSP에 구현하는 것이 가능하여 디지털 이동통신 시스템의 소형화를 통한 비용 절감 효과를 기대할 수 있게 되었다. 따라서 본 논문의 주제는 CS-ACELP 음성 부호화기를 구현하는데 있어서 하나의 DSP로 보다 많은 채널을 서비스할 수 있도록 복잡도(MIPS, ROM, RAM)를 최소화하는 것으로 하였다.

본 논문에서는 Lucent Technologies사의 16 비트 고정 소수점 DSP인 DSP16210을 이용하여 권고안(G.729)과 함께 제공되는 고정 소수점 C 프로그램을 바탕으로 bit-exact 규격에 맞게 DSP16210 기계어 코드로 구현하였다. 또한 구현된 시스템의 음질을 평가하기 위해 임의의 음성 신호에 대해 C 프로그램의 출력과 비교를 하였으며 DSP16210 EVM (Evaluation Module) 보드상에서 절대적(Absolute Comparison Rate, ACR) 음질 평가를 수행 하였다.

본 논문의 2 장에서는 CS-ACELP 음성 부호화기의 기본구조와 특징에 대하여 살펴보고, 3 장에서는 실시간 구현에 대해 논하며, 4 장에서는 구현된 시스템의 음질 시험에 대해 살펴보고, 마지막으로 5 장에서 결론을 맺는다.

### 2. CS-ACELP 음성 부호화기

CS-ACELP는 기본적으로 분석/합성 방식의 CELP[7] 구조를 이용하며 기본 구조는 그림 1과 같다. 8 KHz로

샘플링된 입력신호는 저주파 성분과 오버 플로우(overflow)를 막기위한 전처리 과정을 통과한 후에 10ms의 프레임과 5ms의 부프레임으로 나누어 처리된다. 선형 예측(LPC) 계수들을 추출하기 위한 30ms의 비대칭 윈도우에는 5ms의 새로운 프레임의 데이터가 포함되어 있어서 전체 알고리즘 지연시간은 15ms가 된다.

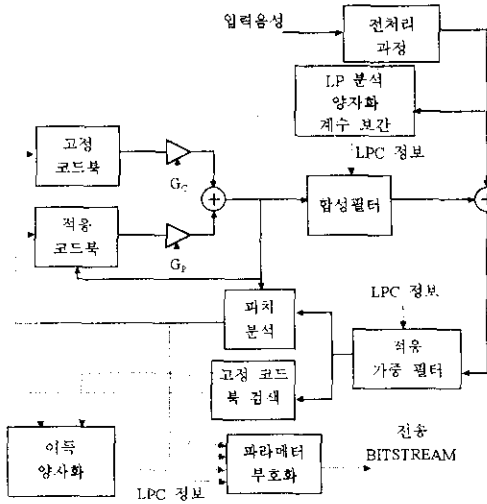


그림 1. CS-ACELP 음성 부호화기의 구조

매 프레임에 대해서 10개의 LPC 계수를 계산하고 이렇게 결정된 LPC 계수는 Line Spectrum Pairs(LSP)로 변환된 후에 2단계의 분리된 벡터 양자화기(VQ)에 의해서 할당된 비트로 양자화 된다. 또한 LSP 양자화 과정에서 에러를 줄이기 위해서 4개의 MA 예측기가 이용되며 각 피치 및 코드북 파라미터를 결정하기 위한 첫번째 부프레임에서의 LSP 값은 바로 전 프레임의 LSP 값과 보간되어 사용된다.

피치 정보를 위한 적응 코드북과 여기 신호를 얻기 위한 고정 코드북 파라미터들은 각 부프레임마다 구해지게 된다. 피치 필터는 적응(adaptive) 코드북의 형태로 사용되며 피치값 결정은 개루프 검색과 페루프 검색을 통해 이루어진다. 부프레임마다 수행되는 페루프 검색은 매 프레임 마다 개루프 검색에서 결정되어진 피치값을 기준으로 1/3의 정밀도를 갖는 비정수형 피치값을 구하는 것이다.

고정 코드북 검색을 위한 target 신호는 피치 검색에서 사용된 target 신호에서 적응 코드북에 의한 기여분을 제거함으로써 얻어진다. 여기신호로 사용되는 대수적 코드북은 표 1과 같이 각 부프레임마다 4개의 펄스만이 지정된 위치에서 +1 또는 -1의 값을 갖고 있다. 이러한 대수적 코드북의 구조에서는 4개의 검색루프를 통해 연속적으로 최적의 펄스 위치를 찾게 되므로 계산량의 감소와 함께 우수한 음질을 얻을 수 있다.

표 1. 8 kbit/s CS-ACELP의 고정 코드북 구조

펄스	부호	위 치
$i_0$	$s_0 : \pm 1$	$m_0 : 0, 5, 10, 15, 20, 25, 30, 35$
$i_1$	$s_1 : \pm 1$	$m_1 : 1, 6, 11, 16, 21, 26, 31, 36$
$i_2$	$s_2 : \pm 1$	$m_2 : 2, 7, 12, 17, 22, 27, 32, 37$
$i_3$	$s_3 : \pm 1$	$m_3 : 3, 8, 13, 18, 23, 28, 33, 38$ 4, 9, 14, 19, 24, 29, 34, 39

적용 코드북과 고정 코드북의 이득값 양자화는 채널 에러에 강한 특성을 갖기 위해 결합구조를 가지는 2개의 벡터 테이블에 의해 양자화 된다. 각각의 테이블은 8개와 16개의 요소값을 가지며 구해진 이득값에 의해 미리 선택된 4개와 8개의 요소값들에 대해서만 검색을 하게 되어 검색시간을 단축시킨다.

CS-ACELP 복호화기(decoder)에서는 전송된 파라미터 값들로부터 복호화기에 사용되어지는 선형예측 계수, 피치, 여기 신호 및 각각의 이득 값을 추출하며 이를 이용하여 합성음을 복원하게 된다. 복원된 합성음의 보다 나은 음질을 위해 후처리(post-processing) 필터를 통과시킨다. 후처리 필터의 계수는 매 부프레임마다 갱신되어지며 long-term, short-term, tilt compensation 필터로 구성되어 있다.

### 3. CS-ACELP의 실시간 구현

본 연구에서 사용된 DSP16210은 Lucent Technologies사의 차세대 기술을 적용한 16비트 고정 소수점 신호 처리 프로세서로 3V의 저전력으로 100 MIPS의 처리 속도와 확장된 캐쉬(cache) 기능을 가지고 있다. 표 2는 DSP16210의 주요 특징을 나타낸 것으로 62 Kword의 내부 메모리가 있어서 시스템을 구성할 때 외부 메모리 별도로 장착하지 않아도 되는 장점이 있으며 워드 및 더블 워드 데이터를 한번에 처리할 수 있는 기능과 2개의 Multiply-Accumulation(MAC)이 있어서 응용 프로그램의 복잡도를 크게 줄일 수 있다.

표 2. DSP16210의 주요 특징

처리 용량	100 MIPS
동작 클럭	50 MHz
동작 전압	3 V
On-chip RAM	62 Kwords
On-chip ROM	8 Kwords
MAC Unit의 수	2 개

CS-ACELP의 실시간 구현은 G.729 표준안과 함께 제공되는 고정 소수점 C 프로그램을 바탕으로 하여 DSP16210 기계어 코드로 구현하였다. C의 각 모듈을 DSP16210 기계어 코드로 작성하고, C의 결과와 구현된 결과가 일치하는지를 검증하기 위해서 테스트 벡터를 C에 인가했을 때 얻어진 입/출력 데이터를 구현된

기계어 코드에 인가하여 결과가 같음을 확인하였다. 위와 같은 방법으로 각 모듈들을 구현한 후, 메인 모듈과 I/O 모듈을 완성하여 그림 2와 같은 소프트웨어 구조를 가지는 프로그램을 완성하였다.

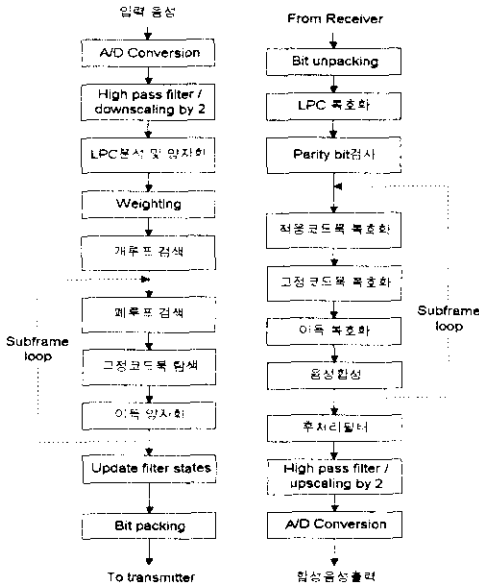


그림 2. CS-ACELP의 소프트웨어 흐름도

CS-ACELP를 구현함에 있어서 DSP16210의 병렬처리 특성과 캐쉬 루프 기능을 최대한 이용하여 복잡도를 줄였다. 표 3은 각 모듈의 계산량을 나타낸 것으로 ACELP 모듈이 상대적으로 많은 계산량이 요구됨을 알 수 있다. 따라서 본 연구에서도 ACELP의 최적화에 많은 노력을 기울였다. 또한 구현된 음성 부호화기는 ITU-T에서 제공되는 모든 테스트 벡터에 대해서 C의 결과와 완전히 일치된 결과를 보였다.

표 3. 구현된 주요 모듈의 계산량

Encoder 전체(13.45 MIPS)	
전 처리부	0.12
LPC 분석 및 양자화	2.37
피치 검색(개루프 검색)	0.98
적용 코드북	1.32
고정 코드북(ACELP)	8.02
이득의 양자화	0.46
기타 모듈	0.28
Decoder 전체(2.09 MIPS)	
파라미터 추출부	0.12
후 처리 필터	1.34
기타 모듈	0.63

하나의 DSP로 다수의 채널을 서비스하는 음성 부호화기는 코드(ROM)를 공유하며 각각의 고유한 데이터(RAM) 영역을 가지고 독립적으로 동작하게 된다. 따라서 태스크들 간의 스케줄링을 위한 간단한 운영체제(OS)가 필수적이며, 실제 시스템에 적용할 때는 반향 제거기(echo canceller)가 포함될 수 있다. 현재 구현되어 있는 반향 제거기는 4 MIPS의 계산량과 0.6 kword의 RAM 및 1 kword의 ROM을 가지며 운영체제의 계산량은 10 MIPS 이내가 될 것으로 추정하고 있다. 따라서 이러한 요소들과 DSP16210의 성능을 고려할 때 구현하고자 하는 CS-ACELP는 표 4에서와 같은 계산량과 메모리 크기가 요구됨을 알 수 있다. 현재 구현된 시스템은 대부분 목표에 부합되는 복잡도를 보여서 4채널을 동시에 처리하는데 충분하며 반향 제거기를 고려하지 않으면 최대 6채널까지 처리할 수 있다.

표 4. 목표 및 구현된 CS-ACELP의 복잡도

	목표치	구현치
계산량(MIPS)	17(under)	15.5
프로그램 메모리(kwords)	15(under)	11.9
데이터 메모리(kwords)	4(under)	4.5
채널 수	4	4

#### 4. 구현된 CS-ACELP의 음질 시험

본 논문에서는 구현된 CS-ACELP 음성 부호화기의 성능을 평가하기 위해 임의의 음성 샘플을 입력하여 C 프로그램의 결과와 구현된 시스템의 결과를 비교하였으며 ITU-T에서 추천하는 Mean Opinion Scores(MOS) 시험방법을 사용하여 구현된 CS-ACELP의 절대적(ACR) 음질평가를 수행하였다.

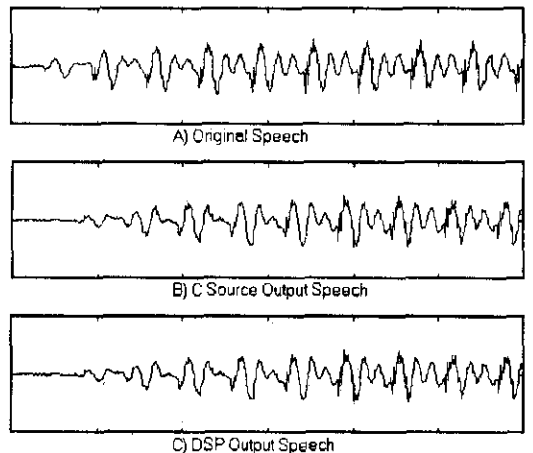


그림 3. 원음, C 및 구현된 시스템의 출력파형 비교

그림 3은 실시간으로 구현된 시스템의 음성 출력

음 원음 및 C 프로그램의 출력과 비교한 것으로 bit-exact 방식으로 개발된 CS-ACELP 음성 부호화기가 C 프로그램과 동일한 결과를 나타냄을 알 수 있다.

DSP16210 EVM 보드상에서 수행된 절대적 음질 평가는 그림 4에서와 같이 입/출력단에 A-law PCM CODEC 을 들으로써 실제 전화망에서의 유사한 환경하에서 실험이 이루어지도록 하였다. 본 실험에서는 잡음이 없는(clean) 상태에서의 음성 신호와 협대역 신호인 DTMF(Dual Tone Multi-Frequency)를 이용하였으며 음질 평가에 사용된 화자는 남자 1명, 여자 1명으로 각 사람은 1 종류의 문장을 말한다. 음질 시험에 참가한 사람은 29 명이며 원음 및 ADPCM 과의 비교에 의해서 MOS 값으로 음질을 평가하게 하였다.

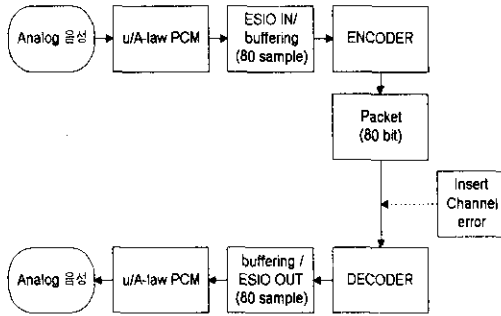


그림 4. 음질 테스트 환경

그림 5에서 나타난 실험결과에서 음성 신호에 대한 절대적 음질 평가에선 일반적으로 알려진 바와 같이 CS-ACELP 음성 부호화기의 음질이 유선 전화의 음질과 유사한 품질을 보였으나, DTMF 에 대해서는 파형 부호화기인 PCM 과 ADPCM 에 비하여 상대적으로 낮은 성능을 보였다. 이러한 비음성 신호에 대한 결과는 보코더가 기본적으로 사람의 음성을 모델링하여 개발되었기 때문인 것으로 판단된다.

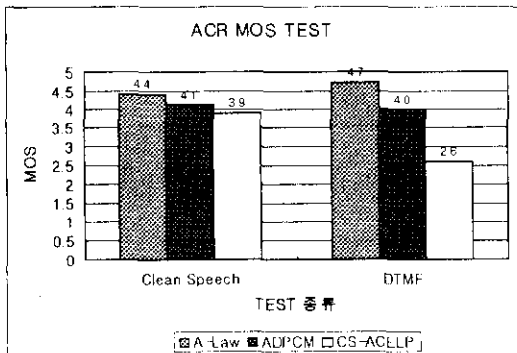


그림 5. 음성과 DTMF 신호에 대한 ACR 결과

## 5. 결 론

본 논문에서는 Lucent Technologies 사의 DSP16210 을 이용하여 CS-ACELP 음성 부호화기를 실시간으로 구현하였다. 구현된 CS-ACELP 는 하나의 신호처리 프로세서로 다수 채널을 서비스할 수 있도록 복잡도(MIPS, ROM, RAM)를 최소화하였으며 ITU-T 에서 제공되는 고정 소수점 C 프로그램을 이용으로 bit-exact 방식으로 구현하였다. 또한 DSP16210 EVM 보드상에 실시간으로 동작되도록 구현되었으며 음성을 스피커를 통해서 직접 듣고 확인할 수 있도록 하여 절대적 음질 평가를 수행 하였다.

현재 구현된 시스템의 계산량은 15.5 MIPS 정도로 DSP16210 의 성능을 고려할 때 반향 제거기, 운영체제 등을 고려하더라도 4 채널 정도를 충분히 처리할 수 있을 것으로 보인다. 앞으로 보다 많은 채널을 하나의 DSP 로 처리하기 위해서는 보다 더 복잡도를 최소화시켜 나가야 하며 작고 효율적인 운영 체제의 개발이 남은 과제이다.

### [참고문헌]

- [1] ITU-T Recommendation G.729, "Coding of speech at 8 kbps using conjugate structure algebraic code excited linear prediction(CS-ACELP)," Geneva, Switzerland, Mar. 1995
- [2] R. V. Cox, "Three New Speech Coders from the ITU Cover a Range of Applications," IEEE Commun. Mag., Sep. 1997
- [3] Mark E. Perkins, Keith Evans and Dominique Pascal, "Characterizing the Subjective Performance of the ITU-T 8 kb/s Speech Coding Algorithm - ITU-T G.729," IEEE Commun. Mag., Sep. 1997
- [4] Simao Ferraz de Campos Neto and Warren Karapetian, "Performance of ITU-T G.729 8 kb/s CS-ACELP Speech Codec with Nonvoice Narrowband Signals," IEEE Commun. Mag., Sep. 1997
- [5] Akitoshi Kataoka, Takehiro Moriya and Shinji Hayashi, "An 8 kb/s Speech Coder Based on Conjugate Structure CELP," Proc. ICASSP '93, pp. 592-595. 1993
- [6] R. Salami, C. Laflamme, J-P. Adoul and D. Massaloux, "A toll quality 8 kb/s speech codec for the personal communication system (PCS)," Proc. IEEE Trans. Veh. Technol., vol. 43, no. 3, pp. 808-816, 1994
- [7] A. M. Kondoz, "Digital Speech," John Wiley & Sons Ltd, 1994
- [8] Lucent Technologies, "DSP16000 Digital Signal Processor Core," Lucent Technologies, Nov., 1997
- [9] Lucent Technologies, "DSP16210 T8502 Codec-DEMO Demonstration Board Version 1.0 A User Manual," Lucent Technologies, April, 1998